Grant agreement for piloting the Framework Partnership Agreement between the National data provider organisations in Denmark and EFSA – Final report

Helwigh, Birgitte; Korsgaard, Helle Bisgaard; Petersen, Pernille Bjørn

Published in:
EFSA Supporting Publications

Link to article, DOI:
10.2903/sp.efsa.2018.EN-1477

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

DTU Food
National Food Institute

EXTERNAL SCIENTIFIC REPORT

APPROVED: 31/07/2018

doi:10.2903/sp.efsa.2018.EN-1477

Grant agreement for piloting the Framework Partnership Agreement between the National data provider organisations in Denmark and EFSA – Final report

Birgitte Helwigh, Helle Korsgaard, Pernille Bjørn Petersen

National Food institute, Technical University of Denmark

Abstract

During the project seven standard operating procedures (SOP) were developed for the four data domains (zoonoses, chemical contaminant, pesticide residues and veterinary medical products residues). The SOPs describe 1) How the national governance of risk assessment data is organised, 2) how data are collected and validated before delivering to EFSA covering all four domains, 3) How data is transferred to EFSA, 4) how to respond to EFSA request for clarification, update and confirmation of data, and 5) how to respond to EFSA request for additional data including PAD requests. Ten enhancements have been developed and implemented in the national reporting system to improve the process of collating and validating data before uploading data to the EFSA DCF. These developments improve the data quality and ensure a higher consistency on the metadata information provided with a data line. The project also included the development of data quality objectives for the four domains. This development was a facilitated by EFSA and carried out as 12 teleconferences.

Key words: Data, Microbiological, Chemical, Contaminants, Pesticides, VMPR, Quality

Question number: EFSA-Q-2018-00568

Correspondence: data.collection@efsa.europa.eu

Acknowledgements: Jytte Butters, Jens-Ole Frimann, Maja Kirkegaard, Katrine Lundsby Jens Hinge Andersen, Søren Johannesen, Helle Lindberg Madsen, Helle Egebjerg Andersen, Maja Effersøe Khan, Anette Nørgaard, Kirsten Halkjær Lund

Suggested citation: National Food Institute, Technical University of Denmark, Birgitte Helwigh, Helle Korsgaard, Pernille Bjørn Petersen, 2018. Grant agreement for piloting the Framework Partnership Agreement between the National data provider organisations in Denmark and EFSA. EFSA supporting publication 2018:EN-1477. 15 pp. doi:10.2903/sp.efsa.2018.EN-1477

ISSN: 2397-8325

Summary

During the project seven standard operating procedures (SOP) were developed for the four data domains (zoonoses, chemical contaminant, pesticide residues and veterinary medical products residues). The SOPs describe 1) How the national governance of risk assessment data is organised, 2) how data are collected and validated before delivering to EFSA covering all four domains, 3) How data is transferred to EFSA and 4) how to respond to EFSA request for clarification, update and confirmation of data, and 5) how to respond to EFSA request for additional data including PAD requests.

Ten enhancements have been developed to improve the process of collating and validating data before uploading data to the EFSA DCF. These developments improve the data quality and ensure a higher consistency on the metadata information provided with a data line. The ten enhancements covered the following areas:

- Automated AMR data load from LIMS to DWH including validation
- Automated selection of AMR isolates in national DWH
- Updating AMR mapping tool
- EFSA catalogues local repository including updating software
- Load platform for excel sheets to DWH
- Updating mapping tools for the zoonoses domain
- Error report implementation on zoonoses data
- Establishing a system to extract the type of analytical method used for the three chemical domains
- Business rule engine for SSD2
- Error reports implementation for the three chemical domains

The project also included the development of data quality objectives for the four domains. This development was a facilitated by EFSA and carried out as 12 teleconferences

Table of contents

1. Introduction

1.1. Background and Terms of Reference as provided by the requestor

Data collected from European Countries are a fundamental component of the EFSA risk assessment process. EFSA has for many years established and supported a process for collecting data from European Competent Authorities data providers with harmonised model and transmission protocol. The EU Member States' competent authorities are by far the major data providers to EFSA.

Since 2007 EFSA awarded to National Competent Authorities involved in data collection 4 M€ to support data harmonisation and electronic transmission of data to EFSA (28 grants and 31 procurements).

These contracts/grants were mainly focused on supporting the implementation of the electronic transmission of data in one specific domain at a time (i.e. zoonoses, occurrence of chemical contaminants and food additives, pesticides residues and veterinary medicinal product residues). No provisions were made in these contracts/grants for maintaining the systems developed in the funded projects. The data submitted by these contracts/grants required in-bound checks of incoming data to EFSA by means of validation rules but did not include any process for measuring, monitoring or rewarding of the data quality. Finally, these contracts/grants were of limited duration. In summary, although in-bound checks of correctness are automatically performed by the EFSA data collection system, the contracts/grants were mainly focusing to achieve the first step of delivering structured data to EFSA rather than to ensuring and improving the quality of the data content and on promoting the systematic resolution of data quality issues.

A need for a more continuous and structured funding from EFSA was highlighted several times by the data providers through the scientific networks, particularly for those data collections where a mandatory legal framework to transmit data to EFSA does not exist. Many countries claim that this prevents the assignment of dedicated resources to continuously deliver high quality data to EFSA. The current contracts are helping to mitigate this issue but only for the duration of the contract, while the major data quality work is needed once the harmonised data collection and transmission system enters the 'production' phase. In addition, EFSA was supporting different data domains with separate contracts. This approach did not encourage the development of data governance at national level.

Therefore, in the context of re-shaping the support to Member States on data collection, an innovative proposal has been developed, with focus to achieve the following benefits:

- Co-ordination and structure at MS level:
  - Having at national level a clear overview on data transmitted to EFSA allowing to guarantee completeness;
  - Establishing data governance;
  - One contact point for data access requests and data licencing.
- Update standards:
  - Ensure stable resources for the data transmission standards updating process;
  - Guarantee updating of national data collection, conversion and transmission systems with the most recent data models and validation rules issued by EFSA.
- Data quality:
  - Dedicated contact point at MS level for data stewarding activities;
  - Improved data quality Performance Indicators (PIs);
  - Capacity for monitoring data quality.

This contract/grant was awarded by EFSA to:

Contractor/Beneficiary: National Food Institute, Technical University of Denmark

Contract/Grant title: Strategic Partnership with Denmark on Data Quality (Pilot project)

Contract/Grant number: GA/EFSA/DATA/2017/01-GA 03

2. Data and Methodologies

2.1. Data

The data used for testing of the enhancements developed during the project are national data on zoonoses, chemical contaminant and pesticide residues sampled in 2016, and veterinary medical products residues (VMPR) sampled in 2017. Since the deadline for reporting of VMPR data has been extended from 30 March to 30 June, if Denmark will not manage to transmit data by the project deadline (30 May) it was agreed that for the successful completion of the project VMPR data sampled in 2016 will be used.

2.2. Methodologies

During this project the National Food Institute (FOOD) has established a data management structure in line with the project description with a national coordinator and data stewards for each of the four domains. In Denmark the responsibility for reporting of data on the zoonoses domain is divided between the data steward and two supporting data stewards. The supporting data stewards are responsible for the collection and validation of data on antimicrobial resistance (AMR) and foodborne outbreaks (FBO), respectively. For the domains on the chemical contaminants, pesticide residuals and VMPR data one steward is responsible for the reporting of data with help from a supervisor and two supporting data stewards. This structure is working well and will continue after the project is finalised. Further, the interactions between the data stewards and the national coordinator will also continue after the project.

For the development of the standard operating procedures (SOP), a process involving the responsible persons for the different steps in the data collection process participated a brain storming meeting where it was discussed which SOPS were relevant, how to divide them into topics and who should be responsible for providing first input on different parts to the national coordinator. The national coordinator wrote up the text and everybody commented on the text before the draft was send to EFSA.

The methodologies used for development of the different enhancements were described in the Technical report on system enhancements (delivery 4 and 10 of the project).

The development of the data quality objectives (DQO) and key performance indicators (KPIs) has been a large part of the project as we participated in all 12 meetings (four meetings for each domain). EFSA was responsible for the process and writing up the DQO and KPI. For each KPI the relevance and challenges for implementation of the KPIs proposed by EFSA in the initial contract was discussed.

3. Assessment/Results

3.1. Activities performed

During the project a total of seven SOPs were written down and approved by EFSA. The SOPs describe 1) How the national governance of risk assessment data is organised, 2) how data are collected and validated before delivering to EFSA covering all four domains, 3) How data is transferred to EFSA and 4) how to respond to EFSA request for clarification, update and confirmation of data. During the process of developing the SOPs the complexity of the collation, validation and reporting of national data to EFSA has become clearer. Meetings together with the main data provider DFVA has further improved the processes and given them a better understanding of the importance of meeting the deadlines in the future.

The development of the 10 enhancements described in the Technical report on system enhancements (delivery 10) has greatly improved the working processes for the relevant areas. In the future more focus can be put on the validation of data rather than entering data into the system as several processes has been automatized.

3.1.1. Activities on quality objectives

The national coordinator and the data stewards have participated actively in all the 12 meetings held during the project. It has been a very long but also fruitful process as the changes implemented into the current version of the DQO and related KPIs has greatly improved the DQOs and the KPIs. Currently we have only seen a pilot version of the calculation of the KPIs for each domain during tele-meetings presented by EFSA, and therefore we have not been able to thoroughly validate the Danish results. We are looking forward to get access to the business intelligence reports on data quality as we believe this will be a good and useful tool in the future. We will provide EFSA with our evaluation of the DQO and KPI once we get the first version of the results.

3.1.2. Proposal on how data quality objectives can be measured and monitored

Some of the DQO and related KPIs are similar for all the domains and it is important that they are measured in the same way and the calculation of the output is the same. We also recommend the numbering of these similar DQO and KPI is the same for all domains. This has been aligned already, but presently we do not have any overview of the results as we do not have access to the calculations of the Danish KPIs.

For all domains some of the KPI has to be adjusted to the data member states report in order to take into account e.g. that changes in the number of flocks tested between years can be larger than the accepted level in the relevant KPI. When the output was presented during the meetings we expressed some doubts that certain low KPIs may not be calculated correctly. We are aware that EFSA worked already to fix some issue, but we would like to see get access to the business intelligence reports on data quality from EFSA on the Danish data, before we can finalise our comments on results or suggest further improvements. However we think the whole idea of developing the DQO and KPIs is good and feasible.

The DQO related to timely Submission should be handled carefully in case of a general postponed deadline (as for the pesticide in 2016-data) or an individually postponed deadline (as for Denmark during reporting of PAH were the update of the catalogue for 'sampMatCode.packmat' as EFSA could not provide the updated catalogue until after the deadline Oct 1st.).

The DQO_ZOO_05 (zoonoses domain) is about the food classification using the MATRIX catalogue, however you should use a different example than Listeria, since the term "ready-to-eat" cannot always be included. However we suggest it should be included in the MATRIX catalogue.

3.2. Achievements, benefits & costs/benefits analysis

3.2.1. Development of Standard Operating Procedures (SOP)

The development of the SOPs has put focus on the process of preparing the national data for reporting to EFSA for all participants in the process as well as leaders. This has proven very useful as all partners understand the chain of work better and the challenges at each link. During the project the discussions on how data is collated for the different domains between the data stewards have provided new ideas for future enhancements. The project has also initiated more collaboration between the stewards across the domains, and between data providers and thus giving a more harmonised reporting.

3.2.2. Output from the enhancements

Enhancement 1: Automated AMR data load from LIMS to DWH including validation

Our system is now able to import an AMR data file from DVFA LIMS, where automated processes harmonise, validate and re-organise the data to fit the FOODs DWH structure. The DWH parameter guided interface ensures that parameters are uniquely defined by IDs, where the corresponding description may be revised at one location for all data in the DWH, as well as presented in Danish or English as selected. With this system, we expect the process of importing, harmonising and revision off LIMS data will be much faster in the future. This means we have better time to make the scientific evaluation of the results, and communicate with the DVFA laboratory regarding unexpected results.

However, we need to figure out procedures when there are errors in the LIMS data, at present we have to manually update the DWH data. The further development of the DWH parameter guided interface is also an important step in the direction of online presentation of the Danish AMR data from animals and meat. The data repository used for creating the XML-files for submission to EFSA may serve as the data base for a web-application. There is a wish for real-time web-based access to the AMR data, however that should be run directly from the DVFA LIMS not on the DWH system at FOOD. If the resources are available at the DVFA, we could develop a system for transferring of data several times a year, and the automated import will enable us to include preliminary annual data online – as well as include 'scanning of critical resistance' at the zoonosis surveillance information meetings with authorities and national stakeholder.

Enhancement 2: Automated selection of AMR isolates in DWH

An automated section function has been developed that selects the same isolates for reporting to EFSA as the original SAS programme. The parameter tables are designed in a way that the selection-criteria may be different between years; in order to accommodate the EU requirements as well as optimise the number of isolates included. For example, there are no EU requirements for Salmonella in caecum samples from pigs, so isolates from a second sample may be included; whereas indicator E. coli isolates from pigs in 2017 may only originate from the first sample. The programming of the automated selection function is based on the sample matrix variables, and summarising results during coding was used to as a validation tool to ensure that the sample matrix hierarchy was harmonised across the different projects. In the future we will include isolates from extra samples from pigs and cattle for Salmonella, Campylobacter and Enterococci. In order to validate the automated selection procedures, we will adapt and run the SAS codes for the 2017 samples

The use of parameter tables has proven very use full in order to harmonise the sample matrix hierarchy. We plan to organise the old data stored in the DANMAP database in the same way.

Enhancement 3: Updating AMR mapping tool

The Danish reporting system is able to submit ESBL/AmpC genes for the isolates that have been analysed by WGS. The DVFAs plan to analyse ESBL/AmpC/CPE producing *E. coli* from the specific monitoring of meat (not the isolates recovered from the specific monitoring of animals).

We should consider including the genes detected from earlier previous surveys in the Data ware house (DWH). However these were recovered by different methodology. We have comparative studies of the two methods, but occurrences were markedly higher when the EURL methodology was used from 2014. These data will not be sent to EFSA, only stored in national DWH for documentation.

Enhancement 4: EFSA catalogues local repository including updating software

All the catalogues are available on a local repository and have made it easier to update our mapping tools. We recommend EFSA also to include the MatrixTool for pesticide residuals on the DCF (XML format).

The new platform for download of catalogues show which catalogues have been updated – this feature is very useful. It would be a valuable improvement if a similar feature could be available for the XML catalogues on the DCF.

Enhancement 5: Load platform for excel sheets to DWH

The development of spread sheets based in the EFSA picklist and the metadata information table has improved the quality of data on Salmonella in poultry, and Histamine and Listeria have ensured a better process for validation of the data by the data provider as the number of handheld steps in the reporting process has been reduced and thus reduced the risk of errors. A new enhancement will develop a system where the data steward will be able to upload the data directly from Excel and Access and the system will create error reports and updates to the data can be carried in tables showing the changes between the original data and the validated data.

Enhancement 6: Updating mapping tools for the zoonoses domain

The development of table with metadata information on data that are automatically uploaded using the platform developed in enhancement 5 has greatly improved the validation of information. The tables can be downloaded into excel and send to the data providers for further validation. This process will enhance the quality assurance and improve the level of information reported to EFSA. If more data are loaded into the DWH at FOOD in the future, the relevant metadata information can be added.

Enhancement 8: Error report implementation on zoonoses data

At the main data provider DVFA, an error report has been developed to address logical errors and warnings for the zoonoses prevalence data. The report has proven an effective tool for visualizing logical error types in microbiological projects. In addition, it can be a valuable tool for pinpointing focus areas in order to prevent the initial typing of errors. Thus, it will be a helpful tool in the future for addressing the task of correcting and minimizing data errors, and thereby enhancing data quality.

Further, this project has initiated a new LEAN project across the DVFA which will take the results of this project further by looking at the procedures for quality control across all data projects in the

DVFA. The purpose is to improve data quality and to establish unified procedures for data quality management across the whole organization.

Enhancement 9: Establishing a system to extract the type of analytical method used for the three chemical domains

The automatic extraction and transfer of method descriptions makes the more manual transfer of lists and test methods and parameters unnecessary and will make the preparatory work in relation to sending data to EFSA more smooth and effective as well as less prone to errors.

The prospects for further use of the automatically transfer system is to create and update mapping tools for the EFSA reporting in our SAS Data manager project (SAS DM project).

Enhancement 10: Business rule engine for SSD2

The SAS files from EFSAs Business rule (BR) engine EFSA was successfully implemented in the data management environment for the domains on chemical contaminants, pesticide residues and VMPR. Some programme modification was implemented in order to integrate the BR-engine in our national environment. Running the national data through the BR-engine gave us error and warning details, so a correction of data could be carried out before uploading to the EFSA DCF and thus ensuring a lower error/warning rate on data and thereby reducing the pressure on EFSA DCF.

To keep performance stability of the engine, it is of outmost importance that EFSA maintain the current project structure of the BR engine in SAS and deliver updated SAS project files with improvements.

The BR-engine had to be integrated into our specific data management environment for the domains on chemical contaminants, pesticide residues and VMPR, and does not work for the zoonoses domain in the current setup. So our BR-engine cannot work as a 'stand-alone' tool to be copied by data stewards from other MS.

Enhancement 11: Error reports implementation on for the three chemical domains

An error report has been developed addressing different logical errors and warnings divided in different chemical domains and projects and a report showing the number of errors corrected. In addition a proposal for a working procedure has been developed and the results are taken further in the organization for a possible implementation across the organization. The report has shown to be an effective tool for visualizing logical error types and will be a helpful tool in the future for addressing the task of correcting and minimizing errors and thereby enhancing data quality.

This project has initiated a new LEAN project across the organization which will take the results of this project further by looking at the procedures for quality control across all data projects in the DVFA. The developed proposal for a working procedure will, if possible, be officially implemented across the organization.

3.3. Resources used at DTU

For the reporting year 2018 we have estimated the number of days spend on the development and implementation of the enhancements, data collation, validation, reporting and upload to EFSA for each domain at DTU Food as well as the number of days used for management and meetings during the pilot FPA project. Unfortunately it is not possible for us to estimate the number of days used during the 2017 reporting for the different domains as our time registrations has not been sufficiently detailed.

Table 1: Resources used at the National Food Institute, Technical University of Denmark on management (general and related to the pilot Framework Partnership Agreement), the development and implementation of the enhancements, data collation, validation, reporting and upload to EFSA from the national repositories to EFSA data warehouse for the reporting year 2018

	Contaminants	Pesticide residues	VMPR	Zoonoses	
				AMR	Prevalence
2018 reporting of 2017 data and development of enhancements (until 31 May 2018)	90 days ≈ Approx. 71,000€	120 days ≈ Approx. 95,000 €	130 days ≈ Approx. 103,000 €	250 days ≈ Approx. 198,000 €	140 days ≈ Approx. 110,000€

Note the resources used on development and implementation of the enhancements, data collation and validation of data by the data provider DVFA are not included

3.4. Challenges

3.4.1. EFSA catalogues, amendments and deadlines

The catalogues are the backbone of the entire reporting system and it is very important that EFSA keep a consistent structure. E.g. terms should not be changed but the STATUS variable should be updated, otherwise new catalogues cannot be used to visualize the historical data in our system.

In most cases, EFSA provide us with a document describing the amendments to the reporting system and this is a very important document for us. E.g. for the pesticide residual, when a new version of both the SSD catalogues file and the MatrixTool for pesticide reporting is delivered, we need the document otherwise it is not possible for us to update our national system.

Further it is very important that the catalogues are not changed during any of the reporting periods, e.g. MatrixTool was changed for the pesticides during the reporting period in 2016.

It is a problem that the amendments and changes to the reporting system is only presented to the MS a few month before the reporting period (e.g. the zoonoses network meeting in November present changes to the 2018 reporting covering 2017 data). This does not give the MS sufficient time to implement the changes before the reporting period and new demands to the metadata information has to be introduced to the data providers already during the project planning phase and not when they are collecting the data for the annual reporting. So amendments presented during the zoonoses network meeting in the fall 2018 should cover samples from 2019 and be reported to EFSA in 2020.

When EFSA postpone their own deadlines e.g. late delivery of templates, updated catalogues or errors in the BR-engines, the Member State deadline should be postponed accordingly. Several times we have experienced that EFSA puts pressure on us to deliver on the initial deadline, but EFSA needs to acknowledge this might not be possible. However, we will always do our outmost to make up for the lost time. KPIs that are influenced by any changes imposed by EFSA should automatically be updated with this information, so they are calculated correct.

3.4.2. Business rules

BR should be advised timely for MS to encounter them in the national reporting system. Any late business rule announcement is not improving data quality. Further, it is essential that EFSA has a

procedure for validating the output from execution of BR's before they are sent to the MS to avoid delay in the processes.

BR should have a unique continuous numbering both general across the four domains and for BR that are specific for a domain; otherwise it will be almost impossible to develop a national system using the BR and the maintenance become impossible. Further, this will ease collaboration between data stewards across the domains.

3.4.3. Validation reports

Validation reports from DCF as well as the validation tables in Microstrategy should be simple and readable. We will suggest discussing the design at a data network meeting.

It is important that the validation reports can be used for validating the submitted data and thereby ensures a high data quality rather than presenting data in a form that can be used in a national summary report. The ackDetails from the DCF has improved considerably for the reporting of 2017 data compared to 2016 data. The creation of frequency tables helps the data stewards to focus on errors and gives a better view of the errors/warnings. This provides a better overview of where to put in efforts to improve data quality.

In the 2016 the validation reports for some domains were created using the SSD1 format; however some MS reported using the SDD2 format. Thus the chemical contaminants data could not be validated by Denmark. It is important that EFSA created the validation reports using the relevant SSD format for each MS when there is an option to use different SSD formats.

There are too many sheets in the validation reports on Microstrategy for the pesticides residual and the zoonoses domains (not including the AMR). Data stewards needs to have an overview of data, which currently is impossible, making it difficult to find corrections point and ensuring data quality.

During the data validation and reporting period it is very important that there is a quick response to questions from EFSA as the work process becomes difficult when data stewards have to wait for a long time before a problem is solved. Fortunately, this is only a problem for some of the domains.

3.4.4. Harmonisation between domains and subdomains

There is a lack of harmonization of catalogues of VMPR, chemical contaminants and pesticides that is causing problem during the reporting of VMPR. In 2016 it was requested by EFSA to double report data covered by the domains. For the zoonoses domain there are also some areas that lack harmonization between the prevalence and AMR reporting in 2017. During the zoonoses network meeting EFSA promised to look into this, however in the manuals this have not been changed. During the final meeting for this project, EFSA explained that BRs for the AMR did not check for the metadata information and we could report the same metadata information for the same samples even though the manual state something else. This should be changed for the 2018 reporting.

4. Conclusions on the pilot and recommendations on future implementation

Overall the pilot project has been a success and has strengthened the collaboration between the national coordinator, data stewards, national data providers and EFSA. The understanding of challenges and demands for each domain has greatly improved between data stewards.

The participation in this project has initiated a new LEAN project across the DVFA. The experiences from the project will form the basis for evaluating the quality control procedures across all data

projects in the DVFA. The purpose is to improve data quality and to establish unified procedures for data quality management across the whole organization.

The development and implementation of the enhancements has improved the data quality and new ideas for future enhancements are already emerging. The commitment to EFSA through a contract that includes funds has been a very strong driver for the departments to allocate sufficient resources needed for the development of the enhancements. We would strongly recommend such a contract for all MS as we are sure the quality and timeliness of reporting will improve in most MS.

If the Framework partnership Agreement is to be implemented in all the MS, then EFSA should reconsider the criteria used for the division of funds between the domains. The resources used on reporting of data across the domains are influenced by several different factors depending on the domain, e.g.:

- The legislative demands for metadata information for each domain
- The number of BR and distribution between errors and warnings
- The complexity of data
- Quality and information flow of data
- Large differences in number of data lines (expanded data vs summary data).
- High output on KPIs for pesticides cover over validation of data 4 times a year as we deliver quarterly national reports
- Result evaluation and action taken for non-compliant samples.

Although the validation of data using the newly developed KPIs has not yet been possible, we believe this will be a very strong and good tool for the future. Once the result become available it is important that a MS can see the records not in compliance for a KPI otherwise it is not possible for MS to validate the output as well as learn from the mistakes. After the first year of implementation of the KPIs it is of outmost importance that EFSA and MS evaluate the outcome. This could be carried out in a working group with members from the relevant scientific data networks and must be discussed thoroughly at scientific network meetings for all four domains.

The use of a BR-engine has had a beneficial influence on the data preparation process as the process is dynamic, depending on different partners (e.g. sampler, laboratory, authorities) and their respond time. Part of the data can be prepared, validated against the BR and corrected while waiting for respond on other parts of the data. The process becomes iterative, more flexible and independent of EFSAs reporting period. For EFSA this will result in less uploads and with fewer errors/warnings once the reporting period starts, thus the pressure on the EFSA DCF will decrease and the waiting time for MS will decrease as well.

We suggest that EFSA develop the BR-engine in a non-proprietary tools/software like R or Python, so it is available to all MS. EFSA should host and maintain this BR-engine and it should function for all four domains. Presently the BR-engine is programmed in SAS, which has suited us perfectly. However, the licence to SAS is fairly costly and not all MSs have access to the programme. The BR-engine should be available online before the reporting season so the data stewards can efficiently scan datasets for errors and missing values during the data preparation process, and reduce the time a submitted XML-file currently are waiting in line to be tested against the BRs at EFSA.

We also suggest a Maximum Residual Level (MRL)-check engine for the Pesticide domain. As for the BR-engine, it should be programmed in non-proprietary tools/software, maintained by EFSA. The engine should be available all year as there is a need for continuously MRL-check of data throughout the year to ease the workload during the reporting period and ensure a higher data quality.

5. Summary

In summary, the most important outputs of the data quality pilot project for Denmark are:

- Increased awareness of data management and how time consuming it is
- Better overview and understanding of the data processes
- Better documentation of data processes and management
- More time allocated to development of data quality enhancements
- Increased awareness on data quality
- Improved data quality
- Increased collaboration between data providers and the data stewards
- Increased collaboration between data domains
- Increased collaboration between data stewards and application developers
- Better compliance to deadlines by data providers

Abbreviations

FOOD	National Food Institute
DVFA	Danish Veterinary and Food Authorities
EFSA	European Food Safety Authority
FBO	Food borne outbreaks
AMR	Antimicrobial resistance
VMPR	Veterinary medical products
SOP	Standard operating procedures
DQO	Data quality objectives
DWH	Data ware house
LIMS	Laboratory information management system
KPI	Key performance indicators
OCC	Occurrence of chemical contaminants
MS	Member States
BR	Business Rules
MRL	Maximum residual level